Available online at www.sciencedirect.com
ScienceDirect
ELSEVIER
Linear Algebra and its Applications 421 (2007) 117–121
LINEAR ALGEBRA AND ITS APPLICATIONS
www.elsevier.com/locate/laa

Weak log-majorization, Mahler measure and polynomial inequalities

Rajesh Pereira *

University of Saskatchewan, Saskatoon, SK, Canada S7N 5E6

Received 19 August 2005; accepted 8 February 2006
Available online 24 April 2006
Submitted by S. Kirkland

Abstract

We explore the use of the weak log-majorization order in the analytic theory of polynomials. We examine the relationship between weak log-majorization and Mahler measure. We also improve the weak log-majorization form of the de Bruijn–Springer–Mahler inequality.
© 2006 Elsevier Inc. All rights reserved.

Keywords: Polynomials; Weak log-majorization; Mahler measure

1. Weak log-majorization and Mahler measure

Whenever roots or eigenvalues are listed in this paper, any root or eigenvalue of multiplicity $m > 1$ will always be listed m times. Any set of real numbers will always be listed in descending order and any set of complex numbers will be listed in descending order of modulus (i.e. $|z_1| \geqslant |z_2| \geqslant \cdots \geqslant |z_n|$).

We begin with a brief review of the weak majorization and weak log-majorization orders.

Definition 1.1. Let $(x_1, x_2, \ldots, x_n)$ and $(y_1, y_2, \ldots, y_n)$ be two n-tuples of real numbers arranged in descending order. Then we say that $(x_1, x_2, \ldots, x_n)$ is weakly majorized by $(y_1, y_2, \ldots, y_n)$ (and we write $(x_1, x_2, \ldots, x_n) \prec_w (y_1, y_2, \ldots, y_n)$) if $\sum_{j=1}^{k} x_j \leqslant \sum_{j=1}^{k} y_j$ for all k; $1 \leqslant k \leqslant n$.

* Fax: +1 306 966 6086.
E-mail address: rjxpereira@yahoo.com

0024-3795/$ - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.laa.2006.02.011

Definition 1.2. Let $(a_1, a_2, \dots, a_n)$ and $(b_1, b_2, \dots, b_n)$ be two n-tuples of non-negative real numbers arranged in descending order. Then we say that $(a_1, a_2, \dots, a_n)$ is weakly log-majorized by $(b_1, b_2, \dots, b_n)$ if $\prod_{j=1}^{k} a_j \leqslant \prod_{j=1}^{k} b_j$ for all k; $1 \leqslant k \leqslant n$.

The following is a useful well-known characterization of weak majorization.

Lemma 1.3 [7, Proposition 4.B.2]. *Let $(x_1, x_2, \dots, x_n)$ and $(y_1, y_2, \dots, y_n)$ be two n-tuples of real numbers. Then $(x_1, x_2, \dots, x_n) \prec_w (y_1, y_2, \dots, y_n)$ if and only if $\sum_{j=1}^{n} \phi(x_j) \leqslant \sum_{j=1}^{n} \phi(y_j)$ for all increasing convex functions ϕ on $\mathbb{R}$.*

We now introduce the concept of Mahler measure which has important applications in number theory and dynamical systems.

Definition 1.4 [4]. Let p be any non-zero complex polynomial, then we define $M(p)$, the Mahler measure of p as follows: $M(p) = \exp\{\int_0^1 \ln|p(\mathrm{e}^{2\pi \mathrm{i} t})| \mathrm{d}t\}$.

Many subsequent researchers prefer to work with the logarithmic Mahler measure defined as $m(p) = \ln(M(p)) = \int_0^1 \ln|p(\mathrm{e}^{2\pi \mathrm{i} t})| \mathrm{d}t$. Using a result of Jensen, Mahler was able to show that $M(p)$ can be expressed in terms of the leading coefficient and roots of p.

Theorem 1.5 [4]. *Let $p(z) = a \prod_{j=1}^{n} (z - z_j)$, then $M(p) = |a| \prod_{j=1}^{n} \max(1, |z_j|)$ (and hence $m(p) = \ln M(p) = \ln(|a|) + \sum_{j=1}^{n} \max(\ln(|z_j|), 0)$).*

Our main result of this section is a relationship between certain Mahler measure inequalities and weak log-majorization.

Theorem 1.6. *Let $p(z) = \prod_{j=1}^{n} (z - a_j)$ and $q(z) = \prod_{j=1}^{n} (z - b_j)$ be two polynomials with the $\{a_j\}_{j=1}^{n}$ and $\{b_j\}_{j=1}^{n}$ listed in descending order of modulus. Then the following are equivalent:*

(1) $\prod_{j=1}^{k} |a_j| \leqslant \prod_{j=1}^{k} |b_j|$ *for all* k; $1 \leqslant k \leqslant n$.
(2) $\sum_{j=1}^{n} \phi(\ln(|a_j|)) \leqslant \sum_{j=1}^{n} \phi(\ln(|b_j|))$ *for all increasing non-negative convex functions on* $\mathbb{R}$ (*with the convention that* $\phi(\ln(0)) = 0$)*.*
(3) $m(p(cz)) \leqslant m(q(cz))$ *for all* $c \in (0, \infty)$.

Proof. Throughout this proof, let $R(p)$ and $R(q)$ be the number of non-zero roots of p and q respectively. If (1) holds, then clearly $R(p) \leqslant R(q)$ and $(\ln(|a_1|), \ln(|a_2|), \dots, \ln(|a_{R(p)}|)) \prec_w (\ln(|b_1|), \ln(|b_2|), \dots, \ln(|b_{R(p)}|))$. Let ϕ be an increasing non-negative convex function on $\mathbb{R}$ (with the convention that $\phi(\ln(0)) = 0$). Then by Lemma 1.3, $\sum_{j=1}^{n} \phi(\ln(|a_j|)) = \sum_{j=1}^{R(p)} \phi(\ln(|a_j|)) \leqslant \sum_{j=1}^{R(p)} \phi(\ln(|b_j|)) \leqslant \sum_{j=1}^{n} \phi(\ln(|b_j|))$. Now suppose (2) holds. Since $m(p(cz)) = \sum_{j=1}^{n} \max(\ln(|a_j|) - \ln(c), 0)$ and the functions $f(x) = \max(x - \ln(c), 0)$ are themselves increasing non-negative convex functions of x, it is clear that (2) implies (3). Let us suppose that $m(p(cz)) \leqslant m(q(cz))$ for all $c \in (0, \infty)$. We note that if $c < \min(|a_{R(p)}|, |b_{R(q)}|)$, then $m(p(cz)) = (\prod_{j=1}^{R(p)} |a_j|) c^{-R(p)}$ and $m(q(cz)) = (\prod_{j=1}^{R(q)} |b_j|) c^{-R(q)}$. Therefore $R(p) \leqslant R(q)$ and if $k > R(q)$, $\prod_{j=1}^{k} |a_j| = 0 = \prod_{j=1}^{k} |b_j|$. So suppose $k \leqslant R(q)$, then $b_k \neq 0$ and

$$
\begin{aligned}
|b_k|^{-k}\prod_{j=1}^{k}|a_j| &\leqslant \prod_{j=1}^{k}\max\left(\frac{|a_j|}{|b_k|},1\right)\\
&\leqslant \prod_{j=1}^{n}\max\left(\frac{|a_j|}{|b_k|},1\right)\\
&= M(p(|b_k|z))\\
&\leqslant M(q(|b_k|z))\\
&= \prod_{j=1}^{n}\max\left(\frac{|b_j|}{|b_k|},1\right)\\
&= |b_k|^{-k}\prod_{j=1}^{k}|b_j|.
\end{aligned}
$$

So (3) implies (1). □

In [5], Mahler proved that for any nth degree polynomial p, we have $M(p'(z)) \leqslant nM(p(z))$. We can restate this as $m(r(cz)) \leqslant m(p(cz))$ for any $c > 0$ where $r(z) = zp'(z)$. Our previous theorem shows that this inequality is equivalent to the following weak log-majorization result of Schmeisser.

Corollary 1.7 [10, Corollary 4]. *Let p be any nth degree polynomial with $n \geqslant 2$. Let $z_1, z_2, \ldots, z_n$ and $w_1, w_2, \ldots, w_{n-1}$ be the roots and critical points of p respectively, ordered so that $|z_1| \geqslant |z_2| \geqslant \cdots \geqslant |z_n|$ and $|w_1| \geqslant |w_2| \geqslant \cdots \geqslant |w_{n-1}|$. Then*

$$\prod_{j=1}^{k}|w_j| \leqslant \prod_{j=1}^{k}|z_j| \quad \forall k : 1 \leqslant k \leqslant n-1.$$

We note that Schmeisser's result can also be derived as a consequence of [6, Corollary 4.4]. We also note that using Theorem 1.6, Mahler's inequality can easily be shown to be equivalent to a similar inequality of de Bruijn and Springer [1, Theorem 12] who proved that if $\{z_j\}_{j=1}^{n}$ and $\{w_j\}_{j=1}^{n-1}$ are the roots and critical points of a polynomial p, then $\sum_{j=1}^{n-1}\phi(\ln|w_j|) \leqslant \sum_{j=1}^{n}\phi(\ln|z_j|)$ for all continuous non-negative increasing convex functions ϕ. Henceforth, we will refer to this result as the de Bruijn–Springer–Mahler inequality.

2. A result of Cheung and Ng

Theorem 2.1. *Let $n \geqslant 2$ and $z_1, z_2, \ldots, z_n \in \mathbb{C}$. Let $p(z) = \prod_{j=1}^{n}(z - z_j)$ and $p_a(z) = \prod_{j=1}^{n}(z - |z_j|)$. Let $w_1, w_2, \ldots, w_{n-1}$ denote the critical points of $p(z)$ listed in descending order of modulus and $v_1, v_2, \ldots, v_{n-1}$ denote the critical points of $p_a(z)$ listed in descending order. Then $\prod_{j=1}^{k}|w_j| \leqslant \prod_{j=1}^{k}v_j$ for $1 \leqslant k \leqslant n-1$.*

This result was proved by Cheung and Ng in the case that one of the roots of p is zero [2, Theorem 2.1]; they asked whether this condition is unnecessary. We will prove that this is indeed the case.

We first need some notation and a few results. For any n by n matrix A, let $\{\lambda_j(A)\}_{j=1}^n$ denote the eigenvalues of A listed in descending order of modulus and let $\{\sigma_j(A)\}_{j=1}^n$ denote the singular values of A listed in descending order. (The singular values of A are the eigenvalues of $|A| = (A^*A)^{\frac{1}{2}}$.)

The first and best known of the log-majorization results is the following result of Weyl.

Lemma 2.2 [13]. *Let A be any n by n matrix. Then $\prod_{j=1}^k |\lambda_j| \leqslant \prod_{j=1}^k \sigma_j$ for $1 \leqslant k \leqslant n-1$ and $\prod_{j=1}^n |\lambda_j| = \prod_{j=1}^n \sigma_j$.*

We will also need the following result of Komarova and Rivin. In what follows, I_n is the n by n identity matrix and J_n is the n by n matrix all of whose entries are ones.

Lemma 2.3 [3, Lemma 5.6]. *Let $n \geqslant 2$ and $z_1, z_2, \ldots, z_n \in \mathbb{C}$. Let $p(z) = \prod_{j=1}^n (z - z_j)$ and $D = \operatorname{diag}(z_1, z_2, \ldots, z_n)$. Then the characteristic polynomial of the matrix $B = D(I_n - \frac{1}{n}J_n)$ is $q(z) = zp'(z)$.*

We can now prove Theorem 2.1 using the essentially same proof as [2, Theorem 2.1] with slightly different matrices.

Proof of Theorem 2.1. Let $D^{\frac{1}{2}} = \operatorname{diag}(\sqrt{z_1}, \sqrt{z_2}, \ldots, \sqrt{z_n})$. (There are two choices for the square root of a non-zero complex number. We will choose the square root which is either a non-negative real number or a complex number with positive imaginary part.) Let $A = D^{\frac{1}{2}}(I_n - \frac{1}{n}J_n)D^{\frac{1}{2}}$, A is similar to $B = D(I_n - \frac{1}{n}J_n)$ so its eigenvalues are $\{w_1, w_2, \ldots, w_{n-1}, 0\}$. Using the same argument as in the proof of [2, Theorem 2.1], it can be shown that $|A| = (D^{\frac{1}{2}})^*(I_n - \frac{1}{n}J_n)D^{\frac{1}{2}}$ which is similar to $|D|(I_n - \frac{1}{n}J_n)$ which has eigenvalues $\{v_1, v_2, \ldots, v_{n-1}, 0\}$. Therefore $\{v_1, v_2, \ldots, v_{n-1}, 0\}$ are also the singular values of A. Theorem 2.1 now follows from Lemma 2.2. □

3. An improved De Bruijn–Springer–Mahler inequality

In this section we will improve the log-majorization (Schmeisser) form of the de Bruijn–Springer–Mahler inequality given in Corollary 1.7.

Along with Theorem 2.1, we will need the following result of Sz.-Nagy [12] which was independently rediscovered by Peyser [8].

Lemma 3.1. *Let p be an nth degree polynomial all of whose roots are real. Let $x_1, x_2, \ldots, x_n$ and $v_1, v_2, \ldots, v_{n-1}$ be the roots and critical points of p listed in descending order. Then $v_j \leqslant (\frac{n-j}{n-j+1}x_j + \frac{1}{n-j+1}x_{j+1})$.*

Applying the above lemma to p_a and then using Theorem 2.1, we obtain:

Theorem 3.2. *Let p be any nth degree polynomial with $n \geqslant 2$. Let $z_1, z_2, \ldots, z_n$ and $w_1, w_2, \ldots, w_{n-1}$ be the roots and critical points of p respectively, ordered so that $|z_1| \geqslant |z_2| \geqslant \cdots \geqslant |z_n|$ and $|w_1| \geqslant |w_2| \geqslant \cdots \geqslant |w_{n-1}|$. Then*

$$\prod_{j=1}^{k} |w_j| \leqslant \prod_{j=1}^{k} \left(\frac{n-j}{n-j+1} |z_j| + \frac{1}{n-j+1} |z_{j+1}| \right) \quad \forall k : 1 \leqslant k \leqslant n-1.$$

Note that this result gives a better bound for $\prod_{j=1}^{k} |w_j|$ than Corollary 1.7. The $k = 1$ case of the Schmeisser form of the de Bruijn–Springer–Mahler inequality is $|w_1| \leqslant |z_1|$ which is a simple consequence of the Gauss–Lucas theorem which states that every critical point of a polynomial is in the convex hull of the roots of that polynomial. It is interesting to note that the $k = 1$ case of Theorem 3.2 is similarly a consequence of the following improved version of the Gauss–Lucas theorem due to Specht. (Both the Gauss–Lucas Theorem and Specht's improvement can be found in Section 2.1 of [9].)

Proposition 3.3 [11]. *Let $n \geqslant 2$, $z_1, z_2, \ldots, z_n$ be complex numbers and $p(z) = \prod_{j=1}^{n} (z - z_j)$. Then all the critical points of p lie in the convex hull of the numbers $\{\frac{n-1}{n} z_j + \frac{1}{n} z_k : 1 \leqslant j, k \leqslant n, j \neq k\}$.*

Acknowledgments

I would like to thank the referee for pointing out the connection between Specht's result and our improved version of the de Bruijn–Springer–Mahler inequality.

References

[1] N.G. de Bruijn, T.A. Springer, On the zeros of composition polynomials, Indag. Math. 9 (1947) 406–414.
[2] W.S. Cheung, T.W. Ng, A companion matrix approach to the study of zeros and critical points of a polynomial, to appear in J. Math. Anal. Appl. (Available online in the articles in press section of J. Math. Anal. Appl. website), in press.
[3] N.L. Komarova, I. Rivin, Harmonic mean, random polynomials and stochastic matrices, Adv. Appl. Math. 31 (2003) 501–526.
[4] K. Mahler, An application of Jensen's formula to polynomials, Mathematika 7 (1960) 98–100.
[5] K. Mahler, On the zeros of the derivative of a polynomial, Proc. Roy. Soc. London Ser. A 264 (1961) 145–154.
[6] S.M. Malamud, Inverse spectral problem for normal matrices and the Gauss–Lucas Theorem, Trans. Am. Math. Soc. 357 (2005) 4043–4064.
[7] A.W. Marshall, I. Olkin, Inequalities: Theory of Majorization and its Applications, Academic Press, New York, 1979.
[8] G. Peyser, On the roots of the derivative of a polynomial with real roots, Am. Math. Monthly 74 (1967) 1102–1104.
[9] Q.I. Rahman, G. Schmeisser, Analytic Theory of Polynomials, Oxford University Press, Oxford, 2002.
[10] G. Schmeisser, Majorization of the critical points of a polynomial by its zeros, Comput. Methods Funct. Theory 3 (2003) 95–103.
[11] W. Specht, Eine Bemerkung zum Satze von Gauss–Lucas, Jahresber. Deutsch. Math.-Verein. 62 (1959) 85–92.
[12] J. Sz.-Nagy, Über algebraische Gleichungen mit lauter reellen Wurzeln, Jahresber. Deutsch. Math.-Verein. 27 (1918) 37–43.
[13] H. Weyl, Inequalities between two kinds of eigenvalues of a linear transformation, Proc. Natl. Acad. Sci. USA 35 (1949) 408–411.